

WEST Search History

[Hide Items](#) | [Restore](#) | [Clear](#) | [Cancel](#)

DATE: Sunday, March 21, 2004

| <u>Hide?</u> | <u>Set Name</u> | <u>Query</u> | <u>Hit Count</u> |
|----------------------------------|-----------------|---|------------------|
| <i>DB=USPT; PLUR=YES; OP=ADJ</i> | | | |
| <input type="checkbox"/> | L4 | l1 and L3 | 13 |
| <input type="checkbox"/> | L3 | 709/224[ccls] | 1666 |
| <input type="checkbox"/> | L2 | L1 same (third party) | 1 |
| <input type="checkbox"/> | L1 | ((request\$ or access\$) near4 ((web page) or (web site))) near8 (profil\$ or characteristic\$ or parameter\$ or statistic\$) | 144 |

END OF SEARCH HISTORY



US006549944B1

(12) **United States Patent**
Weinberg et al.

(10) Patent No.: US 6,549,944 B1
(45) Date of Patent: *Apr. 15, 2003

- (54) USE OF SERVER ACCESS LOGS TO GENERATE SCRIPTS AND SCENARIOS FOR EXERCISING AND EVALUATING PERFORMANCE OF WEB SITES

(75) Inventors: Amir Weinberg, Zoran (IL); Eduardo Alperin, Ramat-Gan (IL)

(73) Assignee: **Mercury Interactive Corporation,
Sunnyvale, CA (US)**

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 108 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: 09/610,909

(22) Filed: Jul. 6, 2000

Related U.S. Application Data

- (63) Continuation of application No. 09/315,795, filed on May 21, 1999, which is a continuation of application No. 08/949,680, filed on Oct. 14, 1997, now Pat. No. 5,974,572, which is a continuation-in-part of application No. 08/840,103, filed on Apr. 11, 1997, now Pat. No. 5,870,559.

(60) Provisional application No. 60/028,474, filed on Oct. 15, 1996.

(51) Int. Cl.⁷ G06F 15/173

(52) U.S. Cl. 709/224; 709/200; 709/203;
714/47; 714/33

(58) Field of Search 709/201, 217,
709/219, 224, 228; 707/10, 12; 705/14;
709/200; 714/1, 25-26, 28; 705/25

References Cited

U.S. PATENT DOCUMENTS

- | | | |
|-------------|-----------|---------------------------------|
| 5,303,166 A | 4/1994 | Amalfitano et al. |
| 5,446,874 A | • 8/1995 | Waclawsky et al. 714/1 |
| 5,537,542 A | • 7/1996 | Eilert et al. 709/201 |
| 5,544,310 A | • 8/1996 | Forman et al. 714/31 |
| 5,657,438 A | 8/1997 | Wygodny et al. |
| 5,696,701 A | • 12/1997 | Burgess et al. 714/25 |
| 5,761,673 A | • 6/1998 | Bookman et al. 707/104 |
| 5,774,123 A | • 6/1998 | Matson 345/357 |
| 5,787,254 A | • 7/1998 | Madalozzo, Jr. et al. . 709/228 |
| 5,812,780 A | 9/1998 | Che et al. |

- | | | | | | |
|-------------|---|---------|-----------------|-------|---------|
| 5,819,066 A | * | 10/1998 | Bromberg et al. | | 707/102 |
| 5,848,415 A | * | 12/1998 | Guck | | 707/10 |
| 5,892,917 A | * | 4/1999 | Myerson | | 709/224 |
| 5,937,390 A | * | 8/1999 | Hyodo | | 705/14 |
| 5,974,572 A | * | 10/1999 | Weinberg et al. | | 714/33 |
| 6,122,740 A | * | 9/2000 | Andersea | | 713/200 |

OTHER PUBLICATIONS

Measuring the Performance of HTTP Daemons by Robert E. McGrath, Feb. 5, 1996 *

(List continued on next page.)

Primary Examiner—Ayaz Sheikh

Assistant Examiner—Thu Ha Nguyen

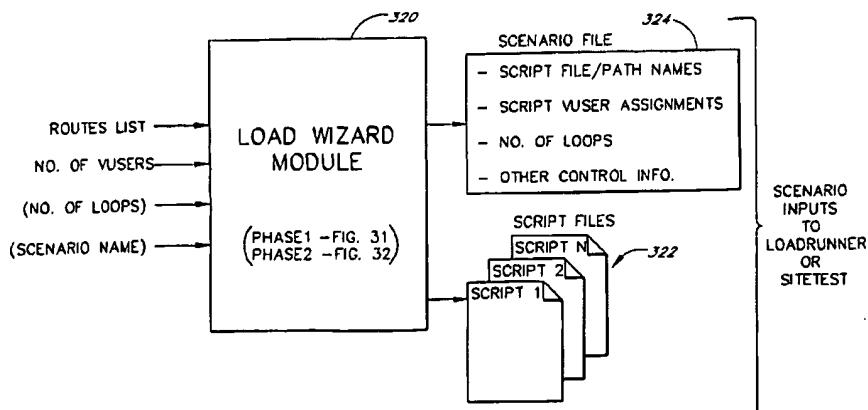
(74) Attorney, Agent, or Firm—Knobbe, Martens, Olson & Bear LLP

ABSTRACT

A visual Web site analysis program, implemented as a collection of software components, provides a variety of features for facilitating the analysis, management and load-testing of Web sites. A mapping component scans a Web site over a network connection and builds a site map which graphically depicts the URLs and links of the site. Site maps are generated using a unique layout and display methodology which allows the user to visualize the overall architecture of the Web site. Various map navigation and URL filtering features are provided to facilitate the task of identifying and repairing common Web site problems, such as links to missing URLs. A dynamic page scan feature enables the user to include dynamically-generated Web pages within the site map by capturing the output of a standard Web browser when a form is submitted by the user, and then automatically resubmitting this output during subsequent mappings of the site. An Action Tracker module detects user activity and behavioral data (link activity levels, common site entry and exit points, etc.) from server log files and then superimposes such data onto the site map. A Load Wizard module uses this activity data to generate testing scenarios for load testing the Web site.

32 Claims, 32 Drawing Sheets

**Microfiche Appendix Included
(1 Microfiche, 51 Pages)**



First Hit Fwd Refs Generate Collection Print

L4: Entry 6 of 13

File: USPT

Apr 15, 2003

DOCUMENT-IDENTIFIER: US 6549944 B1

TITLE: Use of server access logs to generate scripts and scenarios for exercising
and evaluating performance of web sitesDetailed Description Text (136):

During the load testing process, each Vuser monitors the Web site's responses to the client requests submitted by that Vuser, and records various performance-related characteristics of these responses. These characteristics include, for example, response times to individual client requests, timeout events, and error events. Following the load testing process, the user is presented with a set of graphical reports that allow the user evaluate the site's performance.. Using these reports, the user can, for example, compare response times of different site components (Web servers, CGI scripts, APIs, proxy servers, etc.) to identify bottlenecks and other performance problems.

Current US Original Classification (1):

709/224



US006449604B1

(12) **United States Patent**
Hansen et al.

(10) Patent No.: US 6,449,604 B1
(45) Date of Patent: Sep. 10, 2002

- (54) **METHOD FOR CHARACTERIZING AND VISUALIZING PATTERNS OF USAGE OF A WEB SITE BY NETWORK USERS** Web Pages for Webtrends (www.webtrends.com/wt_main.htm).

(75) Inventors: **Mark Henry Hansen, Hoboken; Wlm Sweldens, New Providence, both of NJ (US)** Web Pages for Net.Genesis (www.netgen.com/intro.html).

(73) Assignee: **Lucent Technologies Inc., Murray Hill, NJ (US)** Web Pages for WebManage Technologies (www.webmanage.com/main.html).

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days. Web Pages for Marketwave (www.marketwave.com/frames-default.htm).

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days. Web Pages for WebSell (www likeminds com/solutions/websell/).

(21) Appl. No.: 09/588,488

(22) Filed: Jun. 6, 2000

Related U.S. Application Data

- (62) Division of application No. 09/082,792, filed on May 21, 1998.

(51) Int. Cl.⁷ G06F 17/30; G06F 15/173

(52) U.S. Cl. 707/1; 709/223; 709/224;
709/225

(58) Field of Search 707/1-5, 7, 100,
707/102-104, 513, 526; 709/200-205, 213-220,
224, 225, 226-231, 250

(56) References Cited

U.S. PATENT DOCUMENTS

- | | | |
|----------------|---------|--------------------|
| 5,870,559 A | 2/1999 | Leshem et al. |
| 5,875,446 A | 2/1999 | Brown et al. |
| 5,935,207 A | 8/1999 | Logue et al. |
| 5,958,008 A | 9/1999 | Pogrebinsky et al. |
| 5,974,572 A | 10/1999 | Weinberg et al. |
| 6,182,116 B1 * | 1/2001 | Nammai et al. |
| | | 709/204 |

OTHER PUBLICATIONS

Web Pages for Accrue (www.accrue.com).
Web Pages for Andromedia (www.andromedia.com).
Web Pages for Microsoft Site Server (www.microsoft.com/siteserver/default.asp).

Web Pages for Webtrends (www.webtrends.com/wt-main.htm).

Web Pages for Net.Genesis (www.netgen.com/intro.html).

Web Pages for WebManage Technologies (www.webmanage.com/main.html).

Web Pages for Marketwave (www.marketwave.com/frames-default.htm).

Web Pages for WebSell (www likeminds com/solutions/websell).

* cited by examiner

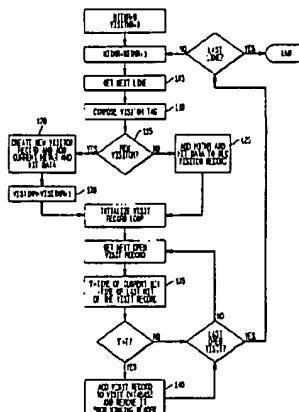
Primary Examiner—Hosain T. Alam

Assistant Examiner—Alford W. Kindred

(57) ABSTRACT

A system is disclosed for displaying information pertaining to the usage of Web pages. The system comprises first and second Web sites. The first Web site comprises plural Web-component files, each having a name in a Web-site directory. The second Web site comprises plural statistics files, each containing usage information about a corresponding Web-component file or sub-directory of Web-component files. The system further comprises a computing device that has a display screen, is operable by a user, and is in communication with the first and second Web sites. The computing device is operated under the control of Web-browser software effective for displaying, on the screen, Web components of the respective Web sites. Significantly, the computing device is effective for requesting and retrieving, from either of the Web sites, data that correspond to user-designated Web components, and it is effective for directing a data request to either of the Web sites in response to user-designation of a Web component from the other Web site.

11 Claims, 6 Drawing Sheets



First Hit Fwd Refs

L4: Entry 7 of 13

File: USPT

Sep 10, 2002

DOCUMENT-IDENTIFIER: US 6449604 B1

TITLE: Method for characterizing and visualizing patterns of usage of a web site by network users

Brief Summary Text (16):

Other software tools provide reports, in the form of HyperText documents, on the usage of selected (such as the most popular) pages. Information from these reports can be displayed via the user-side browser, and links are provided for viewing the selected Web pages. However, these software tools also fail to provide convenient access from a Web page to the statistics that pertain to it.

Detailed Description Text (25):

Hostname: The user who is accessing usage data may wish to filter out his own accesses to the Web site, because they might otherwise skew the statistics. Moreover, filtering on this field may be desirable in order to focus specifically on internal or on external visitors.

Current US Cross Reference Classification (2):

709/224



US006321256B1

(12) United States Patent
Himmel et al.(10) Patent No.: US 6,321,256 B1
(45) Date of Patent: *Nov. 20, 2001(54) METHOD AND APPARATUS FOR
CONTROLLING CLIENT ACCESS TO
DOCUMENTS(75) Inventors: Maria Azua Himmel; Vlktors Berstis,
both of Austin, TX (US)(73) Assignee: International Business Machines
Corporation, Armonk, NY (US)(*) Notice: Subject to any disclaimer, the term of this
patent is extended or adjusted under 35
U.S.C. 154(b) by 0 days.This patent is subject to a terminal dis-
claimer.

(21) Appl. No.: 09/080,018

(22) Filed: May 15, 1998

(51) Int. Cl.⁷ G06F 13/00

(52) U.S. Cl. 709/218; 705/14; 709/224

(58) Field of Search 709/217, 218,
709/219, 224; 705/14(56) References Cited
U.S. PATENT DOCUMENTS

| | | | | |
|-----------|-----------|------------------|-------|---------|
| 5,710,884 | • 1/1998 | Dedrick | | 709/217 |
| 5,754,787 | • 5/1998 | Dedrick | | 709/228 |
| 5,768,521 | • 6/1998 | Dedrick | | 709/224 |
| 5,809,242 | • 9/1998 | Shaw et al. | | 709/217 |
| 5,823,879 | • 10/1998 | Goldberg et al. | | 463/42 |
| 5,848,396 | • 12/1998 | Gerace | | 705/10 |
| 5,850,520 | • 12/1998 | Griebenow et al. | | 709/206 |
| 5,855,008 | • 12/1998 | Goldhaber | | 705/14 |
| 5,991,735 | • 11/1999 | Gerace | | 705/10 |
| 6,009,410 | • 12/1999 | LeMole et al. | | 705/14 |
| 6,119,098 | • 9/2000 | Guyot et al. | | 705/14 |

* cited by examiner

Primary Examiner—Kenneth R. Coulter

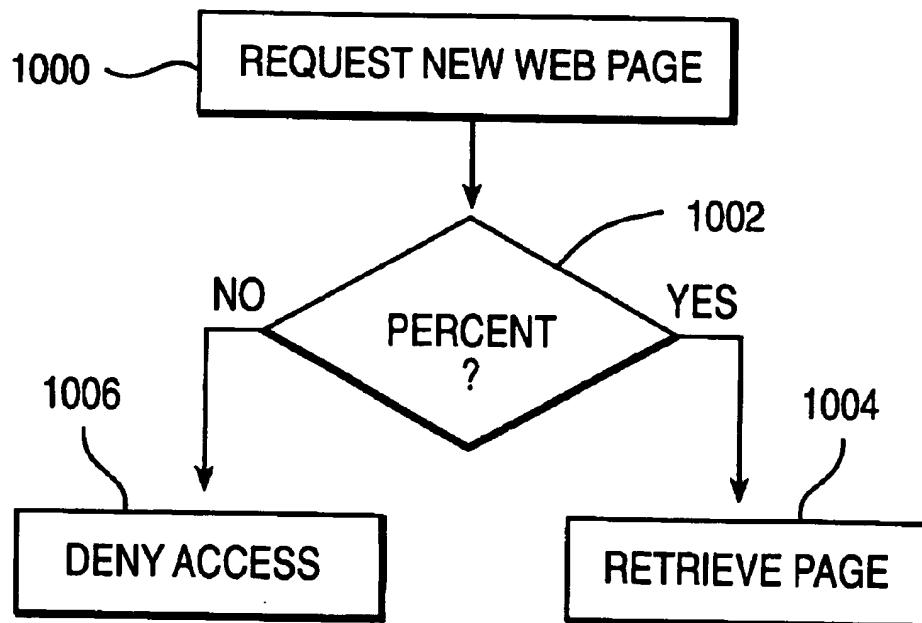
(74) Attorney, Agent, or Firm—Andrea Pair Bryant;
Richard A. Heakler; Leslie A. Van Leeuwen

(57) ABSTRACT

A method and apparatus for detecting, storing and retrieving information, including duration of view time, concerning advertisements included with Web pages seen by a particular user and thereafter using the stored information in controlling access of that user to subsequent Web pages.

3 Claims, 9 Drawing Sheets

X



First Hit Fwd Refs Generate Collection Print

L4: Entry 8 of 13

File: USPT

Nov 20, 2001

DOCUMENT-IDENTIFIER: US 6321256 B1

TITLE: Method and apparatus for controlling client access to documents

Drawing Description Text (12):

FIG. 10 is a flow chart illustrating how the user profile database of FIG. 9 is used to restrict the access of a user to another web page located on the server based upon the number of advertisements viewed on the web page according to a preferred embodiment of the present invention; and

Drawing Description Text (13):

FIG. 11 is a flow chart is shown illustrating the use of the user profile database, and the advertisement database, of FIG. 9 to dynamically customize a requested web page according to prior recorded viewing preferences of a user in accordance with the teachings of the present invention.

Detailed Description Text (51):

Reference now being made to FIG. 10, a flow chart is shown illustrating how the User Profile Database 902 of FIG. 9 is used to restrict the access of a user to another web page located on the server 188 based upon the number of advertisements viewed on the web page 194 according to a preferred embodiment of the present invention. The method begins at step 1000 upon the request by the user of a new web page located on the server 188 via a new URL entry, hyper-link, or the like. Once the server 188 has received this request, it retrieves the prior received and recorded information (e.g. web page 194) to determine whether the user has viewed a pre-determined number of the advertisements A-D 606-612 (Step 1002). If the user has viewed the pre-determined number, then the requested web page is retrieved (Step 1004). If, however, the user has not viewed the pre-determined number, then the user is notified that they must first view X number of advertisements prior to selecting a new page (Step 1006).

Detailed Description Text (52):

Reference now being made to FIG. 11, a flow chart is shown illustrating the use of the user profile database 902, and the Advertisement Database 904, to dynamically customize a requested web page according to prior recorded viewing preferences of a user in accordance with the teachings of the present invention. In the subsequent example, the User Profile Database 902 is used for making substitution/additions of advertisements which have been indicated as of interest to the user based upon the information contained in the User Profile Database 902. This same information could also be used to merely delete those advertisements which have been indicated as not being of particular interest to the user via the User Profile Database 902 as well.

Current US Cross Reference Classification (2):

709/224



US006304904B1

(12) United States Patent
Sathyaranayan et al.

(10) Patent No.: US 6,304,904 B1
(45) Date of Patent: Oct. 16, 2001

(54) METHOD AND APPARATUS FOR COLLECTING PAGE-LEVEL PERFORMANCE STATISTICS FROM A NETWORK DEVICE

(75) Inventors: Seshadri Sathyaranayan; Robert Conrad Knauerhase, both of Portland, OR (US)

(73) Assignee: Intel Corporation, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: 09/002,164

(22) Filed: Dec. 31, 1997

Related U.S. Application Data

(60) Provisional application No. 60/041,003, filed on Mar. 27, 1997.

(51) Int. Cl.⁷ G06F 15/173

(52) U.S. Cl. 709/224; 709/219; 709/222; 709/205; 709/223

(58) Field of Search 395/182.02; 707/102, 707/106, 101, 202, 10; 702/195; 711/133; 709/227, 224, 217, 219, 205, 222, 201, 202, 239, 226, 206, 223; 379/111; 345/335, 329; 370/389; 713/201; 705/114

(56) References Cited

U.S. PATENT DOCUMENTS

| | | | | |
|-----------|---------|---------------|--------------|---------|
| 5,373,375 | 12/1994 | Weldy | 358/523 | |
| 5,485,579 | * | 1/1996 | Hiltz et al. | 709/227 |
| 5,517,612 | 5/1996 | Dwin et al. | 395/166 | |
| 5,544,320 | 8/1996 | Konrad | 395/200.09 | |
| 5,627,886 | * | 5/1997 | Bowman | 379/111 |
| 5,673,322 | 9/1997 | Pepe et al. | 380/49 | |
| 5,684,969 | 11/1997 | Ishida | 395/342 | |
| 5,701,451 | 12/1997 | Rogers et al. | 395/600 | |
| 5,706,434 | 1/1998 | Kremen et al. | 395/200.09 | |
| 5,721,827 | * | 2/1998 | Logan et al. | 709/217 |
| 5,724,556 | 3/1998 | Souder et al. | 395/500 | |
| 5,727,157 | * | 3/1998 | Orr et al. | 709/224 |
| 5,727,159 | 3/1998 | Kikinis | 395/200.76 | |

| | | | | |
|-----------|---------|-------------|-----------------|------------|
| 5,742,905 | 4/1998 | Pepe et al. | 455/461 | |
| 5,751,933 | * | 5/1998 | Dev et al. | 395/182.02 |
| 5,768,510 | 6/1998 | Gish | 395/200.33 | |
| 5,787,471 | * | 7/1998 | Inoue et al. | 711/133 |
| 5,805,735 | 9/1998 | Chen et al. | 382/239 | |
| 5,842,158 | 11/1998 | Plasek | 702/195 | |
| 5,892,917 | * | 9/1999 | Myerson | 709/224 |
| 5,895,470 | * | 4/1999 | Pirolli et al. | 707/102 |
| 5,905,988 | * | 5/1999 | Schwartz et al. | 707/104 |
| 5,931,912 | * | 8/1999 | Wu et al. | 709/224 |
| 5,933,832 | * | 8/1999 | Suzuoka et al. | 707/101 |
| 5,941,944 | * | 8/1999 | Msserly | 709/203 |
| 5,948,061 | * | 9/1999 | Merriman et al. | 709/219 |
| 5,960,409 | * | 9/1999 | Wexler | 705/14 |
| 6,018,619 | * | 1/2000 | Allard et al. | 709/224 |
| 6,112,238 | * | 8/2000 | Boyd et al. | 709/224 |
| 6,141,684 | * | 10/2000 | McDonald et al. | 709/222 |
| 6,240,444 | * | 5/2001 | Fin et al. | 709/205 |

OTHER PUBLICATIONS

Armando Fox and Eric A. Brewer, "Reducing WWW Latency and Bandwidth Requirements by Real-Time Distillation," Fifth International World Wide Web Conference, May 6-10, 1996.

Armando Fox et al., Adapting to Network and Client Variability via On-Demand Dynamic Distillation, University of Cal. at Berkeley, Sep. 1996.

Lucas et al. Statistical Characterization Wide Area IP Traffic IEEE, Sep. 1997.*

* cited by examiner

Primary Examiner—Mehmet B. Geckil

Assistant Examiner—Thong Vu

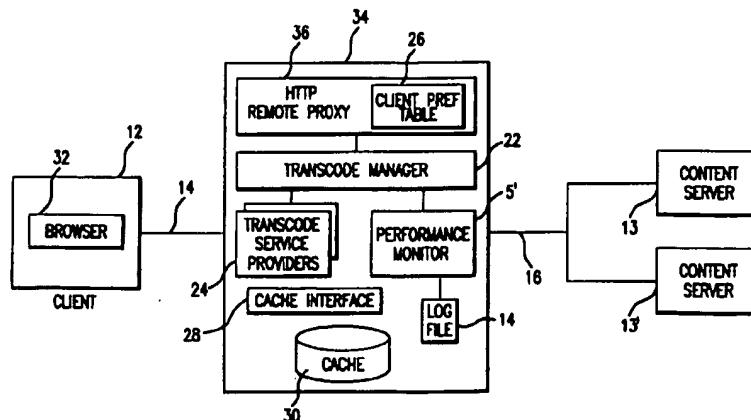
(74) Attorney, Agent, or Firm—Kenyon & Kenyon

(57)

ABSTRACT

A method for collecting statistics from a network device configured to service requests from one or more other devices coupled thereto includes the steps of maintaining a log file containing one or more entries associated with each request serviced by the network device; identifying a page-level request serviced by the network device; and generating statistics associated with the servicing of the page-level request by the network device from the log file entries associated with the page-level request.

9 Claims, 3 Drawing Sheets



[First Hit](#) [Fwd Refs](#)[Search Forms](#)[Generate Collection](#)[Print](#)[Search Results](#)[Help](#)[User Searches](#)

14. ENTRY 9 of 13

File: USPT

Oct 16, 2001

[Preferences](#)[Logout](#)

DOCUMENT-IDENTIFIER: US 6304904 B1

TITLE: Method and apparatus for collecting page-level performance statistics from a network device

Brief Summary Text (6):

In a typical network arrangement for accessing the Internet, a plurality of client devices may be configured to channel requests for Internet resources, such as Web pages, through a network device known as a proxy, or proxy server. For example, proxy servers are often used to channel requests for client devices residing behind a so-called "firewall," or for client devices which use dial-up connections to an Internet service provider (ISP). For a variety of reasons, it may be desirable to collect statistics relating to the performance of such network devices, as well as other devices including content servers. Moreover, it may sometimes be desirable to collect such performance statistics at a page level (that is, compiled with respect to each requested Web page processed by the network device). Unfortunately, there are no existing tools for collecting page-level statistics from such network devices. Instead, known monitoring tools, such as the Webstone.TM. utility distributed by Silicon Graphics Inc., only collect statistical information at a system level. While system-level statistics may be useful for some purposes, for many applications such statistics provide an insufficient level of detail. Accordingly, there is a need for a method and apparatus to collect page-level statistics from a network device.

Current US Original Classification (1):

709/224

CLAIMS:

1. A method for collecting page-level performance statistics from a network device configured to perform transcoding services in connection with responding to requests for web pages by client devices coupled thereto, wherein the requested web pages include one or more associated objects, said method comprising:

servicing a request for a web page by a client device, including retrieving the requested web page and each of its associated objects, transcoding at least one of the retrieved web page or an associated object, and returning the web page and its associated objects to the client device;

maintaining a log file containing a plurality of entries associated with each request for a web page serviced by the network device, the plurality of entries comprising a page-level entry corresponding to the web page and one or more object-level entries corresponding to the objects associated with the web page;

identifying a page-level entry in the log file for a given web page request serviced by the network device;

identifying each object-level entry in the log file for objects associated with the web page; and

generating page-level performance statistics associated with the servicing of the



US006185614B1

(12) **United States Patent**
Cuomo et al.

(10) **Patent No.:** US 6,185,614 B1
(b) **Date of Patent:** Feb. 6, 2001

(54) **METHOD AND SYSTEM FOR COLLECTING USER PROFILE INFORMATION OVER THE WORLD-WIDE WEB IN THE PRESENCE OF DYNAMIC CONTENT USING DOCUMENT COMPARATORS**

(75) Inventors: Gennaro A. Cuomo, Apex; Binh Q. Nguyen, Cary; Sandeep K. Singhal, Raleigh, all of NC (US)

(73) Assignee: International Business Machines Corp., Armonk, NY (US)

(*) Notice: Under 35 U.S.C. 154(b), the term of this patent shall be extended for 0 days.

(21) Appl. No.: 09/084,452

(22) Filed: May 26, 1998

(51) Int. Cl.⁷ G06F 15/173; G06F 15/16;
G06F 7/00

(52) U.S. Cl. 709/224; 709/203; 707/104

(58) Field of Search 709/203, 224;
707/6, 10, 104, 501, 513, 3, 5

(56) **References Cited**

U.S. PATENT DOCUMENTS

- | | | | | |
|-------------|---------|------------------|-------|---------|
| 5,649,186 * | 7/1997 | Ferguson | | 707/10 |
| 5,732,218 | 3/1998 | Bland | | 709/204 |
| 5,740,430 * | 4/1998 | Rosenberg et al. | | 707/200 |
| 5,745,900 * | 4/1998 | Burrows | | 707/102 |
| 5,813,007 * | 9/1998 | Nielsen | | 707/10 |
| 5,890,164 * | 3/1999 | Nielsen | | 707/201 |
| 5,892,917 | 4/1999 | Myerson | | 709/204 |
| 5,893,908 * | 4/1999 | Cullen et al. | | 707/5 |
| 5,895,470 * | 4/1999 | Pirolli et al. | | 707/102 |
| 5,898,836 * | 4/1999 | Freivald et al. | | 709/218 |
| 5,909,677 * | 6/1999 | Broder et al. | | 707/3 |
| 5,913,208 * | 6/1999 | Brown et al. | | 707/3 |
| 5,941,944 * | 8/1999 | Messerly | | 709/203 |
| 5,978,842 * | 11/1999 | Noble et al. | | 709/218 |
| 5,983,268 * | 11/1999 | Freivald et al. | | 709/218 |
| 5,987,480 * | 11/1999 | Donohue et al. | | 707/501 |
| 5,999,929 * | 12/1999 | Goodman | | 707/7 |

6,012,087 * 1/2000 Freivald et al. 709/218

FOREIGN PATENT DOCUMENTS

9831155 7/1998 (WO).

OTHER PUBLICATIONS

Brin, S., et al., "Copy Detection Mechanisms for Digital Documents," Proc. of the 1995 ACM SIGMOD Int'l. Conf. on Management of Data, ACM, pp. 398-409, May, 1995.*
Garcia-Molina, H., et al, "dSCAM: Finding Document Copies Across Multiple Databases," Proc. of the 4th Int'l. Conf. on Parallel and Distributed Information Systems, IEEE, pp. 68-79, May 1995.*

* cited by examiner

Primary Examiner—Ahmad F. Matar

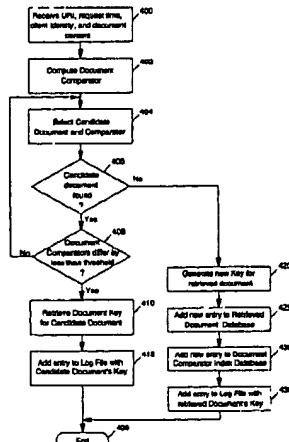
Assistant Examiner—Andrew Caldwell

(74) Attorney, Agent, or Firm—A. Bruce Clay

(57) **ABSTRACT**

Disclosed is a method and system for collecting profile information about users accessing dynamically generated content from one or more servers. In a specific embodiment, a server dynamically generates a web page in response to a user request. The server customizes the web page content based on the requested universal resource identifier (URI) and one or more of: the user's identity, access permissions, demographic information, and previous behavior at the site. The web server then passes the URI, user identity, and dynamically generated web page to an access information collector. The access information collector generates document comparators from the current web page content and compares them to document comparators associated with previously retrieved web pages. If the current web page is sufficiently similar to some previously retrieved web page, the access information collector logs the URI, user identity, and a document key associated with the matching previously retrieved page. Otherwise, the access information collector generates a new key; stores the new key and the document comparators in a database; and logs the URI, user identity, and the newly generated document key.

27 Claims, 4 Drawing Sheets



First Hit Fwd Refs

L4: Entry 10 of 13

File: USPT

Feb 6, 2001

DOCUMENT-IDENTIFIER: US 6185614 B1

TITLE: Method and system for collecting user profile information over the world-wide web in the presence of dynamic content using document comparators

Brief Summary Text (2):

This invention relates in general to computer software, and in particular to a method and system for collecting profile information about users accessing Web pages from a plurality of Web servers. More particularly, the present invention relates to a method and system by which user profile information can be collected when the Web content is generated dynamically for each request at the Web server.

Current US Original Classification (1):

709/224



US005892917A

United States Patent

[19]

Myerson**Patent Number:** 5,892,917**[45] Date of Patent:** Apr. 6, 1999

[54] **SYSTEM FOR LOG RECORD AND LOG EXPANSION WITH INSERTED LOG RECORDS REPRESENTING OBJECT REQUEST FOR SPECIFIED OBJECT CORRESPONDING TO CACHED OBJECT COPIES**

[75] Inventor: Terry J. Myerson, Sunnyvale, Calif.

[73] Assignee: Microsoft Corporation, Redmond, Wash.

[21] Appl. No.: 534,488

[22] Filed: Sep. 27, 1995

[51] Int. Cl.⁶ G06F 7/02

[52] U.S. Cl. 395/200.54; 395/838

[58] Field of Search 340/825.33; 379/34; 380/4; 395/184.01, 651, 200.54, 200.53, 838, 892

[56] **References Cited**

U.S. PATENT DOCUMENTS

- | | | | | |
|-----------|---------|------------------------|-------|------------|
| 3,727,186 | 4/1973 | Stephenson, Jr. et al. | | 340/825.33 |
| 4,220,771 | 9/1980 | Kraushaar et al. | | 379/138 |
| 4,559,416 | 12/1985 | Theis et al. | | 379/34 |
| 5,023,907 | 6/1991 | Johnson et al. | | 380/4 |
| 5,051,944 | 9/1991 | Fetterolf et al. | | 395/184.01 |
| 5,428,785 | 6/1995 | Morel et al. | | 395/651 |
| 5,675,510 | 10/1997 | Coffey et al. | | 395/200.54 |

Primary Examiner—Thomas C. Lee

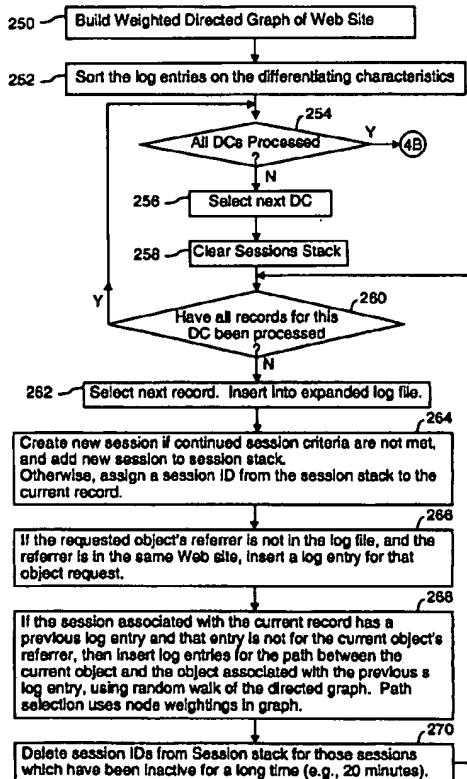
Assistant Examiner—Anderson I. Chen

Attorney, Agent, or Firm—Lee & Hayes, PLLC

[57] **ABSTRACT**

A system and method for analyzing a Web site log file and generating an expanded log file that compensates for information caching and gateway based Web site access. More particularly, the log file expansion procedure of the present invention works with a log file stored in memory on the server computer. The log file contains a sequence of log records, each log record representing an object request by a client computer. Each log record includes data identifying the requested object as well as some data, such as an Internet address, associated with the client computer or a gateway through which the client computer requested the object. The log expansion procedure analyzes the sequence of log records so as to detect object request patterns indicating that object requests not represented by the log records were satisfied by cached object copies, and then supplements the sequence of log records with inserted log records representing object requests for the objects corresponding to the cached object copies. As a result, the supplemented sequence of log records more accurately represents object requests made by client computers than the initial sequence of log records in the log file. Usage metering and analysis procedures utilized the supplemented sequence of log records to generate analysis reports indicative of object request patterns by the client computers.

34 Claims, 6 Drawing Sheets



First Hit Fwd Refs**End of Result Set** Generate Collection

L4: Entry 13 of 13

File: USPT

Apr 6, 1999

DOCUMENT-IDENTIFIER: US 5892917 A

TITLE: System for log record and log expansion with inserted log records representing object request for specified object corresponding to cached object copies

Brief Summary Text (11):

It is a further object of the present invention to assign object requests in the expanded log files to synthesized client sessions so as to represent, in a statistically accurate manner, the number of client sessions accessing a Web site and the distribution of objects accessed by those client sessions.

Brief Summary Text (12):

Another object of the present invention Is to generate analyses of Web site usage based on an expanded log file that represents in a statistically accurate manner the information access patterns of the clients of the Web site.

Brief Summary Text (14):

In summary, the present invention is a system and method for analyzing a Web site log file and generating an expanded log file that compensates for information caching and gateway based Web site access. The expanded log file represents in a statistically accurate manner the information access patterns of the clients of the Web site, although the individual synthesized client sessions represented by the expanded log file do not necessarily represent actual client sessions.

Current US Original Classification (1):

709/224

L4: Entry 13 of 13

File: USPT

Apr 6, 1999

DOCUMENT-IDENTIFIER: US 5892917 A

TITLE: System for log record and log expansion with inserted log records representing object request for specified object corresponding to cached object copies

Brief Summary Text (11):

It is a further object of the present invention to assign object requests in the expanded log files to synthesized client sessions so as to represent, in a statistically accurate manner, the number of client sessions accessing a Web site and the distribution of objects accessed by those client sessions.

Brief Summary Text (12):

Another object of the present invention Is to generate analyses of Web site usage based on an expanded log file that represents in a statistically accurate manner the information access patterns of the clients of the Web site.

Brief Summary Text (14):

In summary, the present invention is a system and method for analyzing a Web site

log file and generating an expanded log file that compensates for information caching and gateway based Web site access. The expanded log file represents in a statistically accurate manner the information access patterns of the clients of the Web site, although the individual synthesized client sessions represented by the expanded log file do not necessarily represent actual client sessions.

Current US Original Classification (1):

709/224